Acta Psychologica 144 (2013) 97–103

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

Expert intuitions: How to model the decision strategies of airport customs officers?☆

Thorsten Pachur [a,*], Gianmarco Marinello [b]

[a] Max Planck Institute for Human Development, Berlin, Germany
[b] Department of Psychology, University of Basel, Switzerland

ARTICLE INFO

Article history:
Received 2 November 2012
Received in revised form 10 May 2013
Accepted 11 May 2013
Available online xxxx

PsycINFO classification:
2340
3430

Keywords:
Decision strategies
Expertise
Heuristics
Customs
Strategy selection

ABSTRACT

How does expertise impact the selection of decision strategies? We asked airport customs officers and a novice control group to decide which passengers (described on several cue dimensions) they would submit to a search. Additionally, participants estimated the validities of the different cues. Then we modeled the decisions using compensatory strategies, which integrate many cues, and a noncompensatory heuristic, which relies on one-reason decision making. The majority of the customs officers were best described by the noncompensatory heuristic, whereas the majority of the novices were best described by a compensatory strategy. We also found that the experts' subjective cue validity estimates showed a higher dispersion across the cues and that differences in cue dispersion partially mediated differences in strategy use between experts and novices. Our results suggest that experts often rely on one-reason decision making and that expert–novice differences in strategy selection may reflect a response to the internal representation of the environment.

1. Introduction

How does practice and experience with a task affect decision making? The expertise literature offers two opposing views on this question.[1] On the one hand, it is commonly assumed that "skill is … often an ability to deal with large amounts of information quickly and efficiently" (Kahneman, 2011, p. 458). As a consequence, experts, by virtue of their extensive familiarity with a domain, can rely on pattern matching, whereas novices have to process the information in a more piecemeal fashion (e.g., Chase & Simon, 1973; Gobet & Simon, 1996; Klein, 1998). Accordingly, it has been argued that expert decision making can be well described by models that integrate multiple pieces of diagnostic information (i.e., cues) in a compensatory—and maybe automatic—fashion (e.g., Glöckner, Heinen, Johnson, & Raab, 2012; Phelps & Shanteau, 1978; see also Glöckner & Betsch, 2008).

On the other hand, some findings indicate the opposite. Specifically, it has been shown that experts consider less information than novices, possibly because they are better able to distinguish relevant from irrelevant cues and also have more knowledge about intercorrelations between cues (Shanteau, 1992). As a consequence, compared to novices experts may be more likely to rely on simple strategies that exploit cue hierarchies, have stopping rules, and thus can lead to noncompensatory decisions. Initial evidence that experts rely on noncompensatory decision strategies was provided by Garcia-Retamero and Dhami (2009). They asked two groups with considerable expertise of house security (experienced burglars and police officers) and a novice group (students) to judge which of two residential properties is more likely to be burgled. It emerged that the decisions of the two expert groups were best modeled by a simple noncompensatory heuristic, whereas the decisions of the novice group were best modeled by a compensatory strategy.

Despite this evidence for experts' reliance on simple strategies, it is currently unclear how generally this result holds. Does it generalize to domains in which experts make hundreds of decisions every day, often receive explicit instructions on which cues to use, and receive regular feedback on the quality of their decisions? Ettenson, Shanteau, and Krogstad (1987) found no difference between experts and novices in terms of the number of cues used in auditing decisions, which the experts made on a daily basis. Our goal in this article is to model the decision strategies of airport customs officers who screen passengers for items such as illegal substances and dutiable goods. Customs officers' decisions have substantial consequences. In 2010, for instance,

☆ We thank Laura Wiles and Susannah Goss for editing the manuscript.
* Corresponding author at: Max Planck Institute for Human Development, Center for Adaptive Rationality, Lentzeallee 94, 14195 Berlin, Germany. Tel.: +49 30 82406335.
E-mail address: pachur@mpib-berlin.mpg.de (T. Pachur).
[1] Consistent with Camerer and Johnson (1991), we define an expert as "a person who is experienced at making predictions in a domain and has some professional or social credentials" (p. 196). That is, our definition does not presuppose that an "expert" necessarily makes better decisions than a novice. This approach is common and has proven useful in the decision sciences, where there have been several demonstrations of alleged "experts" showing unexpectedly poor judgment performance (e.g., Meehl, 1954). Note, however, that in other domains (e.g., music, sports), experts are often defined on the basis of their performance (e.g., Ericsson & Charness, 1994).

0001-6918/$ – see front matter
http://dx.doi.org/10.1016/j.actpsy.2013.05.003

smuggled goods confiscated at borders of the European Union totaled more than €1 billion (European Commission Taxation and Customs Union, 2012). Nevertheless, there is practically no systematic research on the cues and cognitive mechanisms underlying customs officers' decisions. In addition, we investigate a potential mediator of expert–novice differences in strategy use. Specifically, given that the dispersion of cue validities (which indicate how diagnostic a cue is) has been found to be an important factor in strategy selection (Mata, Schooler, & Rieskamp, 2007), we examine whether experts and novices have different internal representations of the cue validity distribution and whether such differences mediate differences in strategy use. Because of their limited experience, novices might be more conservative in their cue validity estimates, resulting in a less disperse distribution; in such an environment, the use of a compensatory strategy is more appropriate (Hogarth & Karelaia, 2007).

Finally, we address a possible confound in the study by Garcia-Retamero and Dhami (2009). Their novice group was both considerably younger and better educated than their expert group and also had a rather different gender distribution. Given that, for instance, age has been shown to be associated with an increased tendency to rely on simple strategies (for an overview, see Mata et al., 2012), these confounds may compromise the conclusion that the observed differences in strategy use were due to differences in expertise.[2] In our study, the control group was closely matched to the expert group in terms of age, education, and gender.

We model expert and novice decisions using two prominent classes of decision strategies. The first class comprises the *weighted-additive strategy* (Payne, Bettman, & Johnson, 1993), according to which the object with the higher sum of positive cue values, multiplied by their respective cue weights, is chosen; and the *equal-weight strategy* (Dawes, 1979), according to which the object with the higher sum of positive cue values, with all cues weighted equally, is chosen. Both strategies are compensatory in nature and represent straightforward implementations of the notion that decisions involve the evaluation of multiple cues. According to the view that experts often rely on (automatic) processing of entire cue patterns (e.g., Klein, 1998), their decisions should thus be best described by these compensatory strategies. The second class of decision strategies is represented by the *take-the-best heuristic* (Gigerenzer & Goldstein, 1996), a noncompensatory, lexicographic strategy with limited search that relies on a single cue to make a decision. Specifically, take-the-best inspects cues sequentially in descending order of validity and compares the alternatives on these cues; inspection of cues is stopped as soon as the alternatives differ on a given cue, and the alternative with a positive value on that cue is chosen (for an investigation of the neural underpinnings of using take-the-best, see Khader et al., 2011). According to the view that experts often rely on decision processes that exploit cue hierarchies (Garcia-Retamero & Dhami, 2009), their decisions should be best described by a noncompensatory heuristic. Because existing empirical evidence for take-the-best is almost exclusively based on artificial laboratory tasks (for an overview, see Bröder, 2011), support for use of take-the-best in the domain of customs decisions would be an important demonstration that it can also describe decision making in applied and more realistic settings (Lipshitz, 2000).

In addition to elucidating possible differences in the cognitive mechanisms underlying expert and novice decision making, our study also contributes to a better understanding of the factors influencing strategy selection. The idea that people have a repertoire of strategies at their disposal has been criticized to the extent that it requires an additional mechanism for selecting how to decide (e.g., Newell, 2005). Several potential mechanisms underlying strategy selection have been proposed, such as a deliberate trade-off between the costs and benefits of a strategy (Payne et al., 1993), reinforcement processes (Rieskamp & Otto, 2006), and factors arising from an interplay between mind and environment (Marewski & Schooler, 2011). To evaluate these proposals, one important step is to map the boundary conditions of people's use of different strategies—and expertise may be one such condition.

2. Method

2.1. Participants

For the group of experts, we recruited 31 customs officers (mean age: 46.9 years, range 32–60) from two international airports in Switzerland (Zürich and Basel). They were mostly male (28; 90.3%) and had professional experience as customs officers of, on average, 15.7 years ($SD = 10.1$, range 2–39). All customs officers indicated their highest educational attainment to be "Sekundarschule" (comparable to high school) with additional vocational training. The novice group consisted of 40 participants matched to the experts in terms of gender, age, and education: they were mostly male (36; 90%), had a mean age of 48.0 years (range 31–61), and like the experts, indicated their highest educational attainment to be "Sekundarschule" with additional vocational training. All participants received 10 Swiss francs as compensation.

2.2. Material and design

Based on interviews with the chief customs officer at a major international Swiss airport, we first identified eight passenger characteristics that are potentially diagnostic for passengers smuggling drugs. These characteristics were (in decreasing importance according to the chief officer): flight origin, gender, nationality, age, amount of luggage, eye contact with officer, clothes, and speed of gait. Moreover, we identified which values on these cues would be more or less indicative of a person smuggling drugs. The resulting positive and negative cue values (indicating whether a person would be more or less likely to smuggle drugs, respectively) for each cue are shown in Table 1.

We then created pairs of passenger profiles (consisting of positive and negative cue values) that would allow us to discriminate between the use of a compensatory versus a noncompensatory strategy. In a first step, we selected from the total of $2^8 \times (2^8 - 1) / 2$ possible pairs of profiles with 8 binary cues those pairs for which take-the-best would make different predictions from both weighted-additive (using linearly decreasing weights from .9, .8, etc., to .1) and equal-weight, excluding cases where a strategy had to guess. In a second step, we sampled randomly from the remaining pairs three sets of 15 pairs each for which take-the-best discriminated on the first, second, and third cue, respectively. This yielded a total of 45 pair comparisons.

Finally, we assigned the cues in the profiles to the verbal cues in Table 1 according to the cue ranking of the chief customs officer. For instance, a cue profile [1, 0, 0, 1, 1, 0, 1, 1] (with cues in descending order of importance) translated into the following passenger profile:

Table 1
Positive (i.e., indicating a greater chance that the person is smuggling drugs) and negative (i.e., indicating a lower chance that the person is smuggling drugs) values for the eight cues.

Cue	Positive value	Negative value
Flight origin	South America	Europe
Gender	Male	Female
Nationality	Domestic	Nondomestic
Age	20–40	40–70
Amount of luggage	One bag	Several bags
Eye contact with officer	No	Yes
Clothes	Casual	Business
Gait	Hurried	Normal

[2] Garcia-Retamero and Dhami (2009) reported that within the expert and novice groups, age and education were uncorrelated with strategy use. However, given that age and education were rather homogenous within the groups, these correlations are possibly restricted due to limited variability.

flying in from South America, female, nondomestic origin, aged 20–40, with one bag, having eye contact with the customs officer, in casual clothes, and with a hurried gait.

We asked participants to place themselves in the position of a customs officer at an international airport. They were presented with a total of four tasks. In a *decision task*, they were shown the 45 pairs of passenger profiles (see Fig. 1) and asked to indicate for each pair which of the two passengers they thought would be more likely to smuggle illegal drugs. The next task was a *ranking task*, in which participants were asked to rank the eight cues in terms of their validity for judging whether a passenger is smuggling drugs. In a *cue validity estimation task*, participants were presented with the cues in the order they had indicated in the ranking task and asked to provide a continuous rating for how diagnostic each cue was for identifying a drug smuggling passenger, using a scale from 1 (= not diagnostic) to 100 (= highly diagnostic). (We used each participant's responses in the ranking task and the cue validity estimation task when deriving the predictions of the different decision strategies; see below for details.) Finally, in a *confidence task*, participants indicated their subjective confidence in the accuracy of their validity ranking for each cue using a scale from 1 (= absolutely uncertain) to 7 (= absolutely certain). In addition, all participants provided demographic information, and the officers also indicated the number of years they had been working for the customs (i.e., their professional experience).

2.3. Procedure

All tasks were administered on a computer, presented in the order indicated above. The order in which cues were arranged on the screen was determined randomly in each trial. The tasks were self-paced and participants took, on average, around 20 min to complete all tasks.

3. Results

3.1. Did experts and novices differ in their cue validity estimates?

In ranking the cues according to their validity, the experts showed a higher consensus than the novices, Kendall's $W = .55$ ($p = .001$) vs. .16 ($p = .001$). For instance, 30 of the 31 officers, but only 19 of the 40 novices estimated the "flight origin" cue to be the most important one. Fig. 2 shows the mean validity estimate for each of the 8 cues. Across the cues, the ordinal agreement between the experts' and the novices' estimates was rather low, $r_s = .17$ ($p = .69$). The experts' ranking agreed more closely with the chief officer's (see Method and Table 1) than did the novices' ranking ($r_s = .43$, $p = .29$, vs. $r_s = .14$, $p = .74$). A mixed-design ANOVA with participants' cue validity estimates as dependent variable and group (experts vs. novices) and cue (Cues 1–8) as independent variable (with the latter being a within-subjects factor) showed no main effect of expertise, $F(1,\ 69) = 1.4$, $p = .24$, but both a main effect of cue, $F(5.26,\ 363.13) = 28.9$, $p = .001$, and an interaction between cue and group, $F(5.26,\ 363.13) = 15.6$, $p = .001$. This indicates that the experts and the novices differed in how they evaluated the individual cues. In particular, the novices estimated the validity of the "gait" and "eye contact" cues to be substantially higher than did the experts, whereas the experts estimated the validity of the "flight origin", "nationality", and "luggage" cues to be substantially higher than did the novices (Fig. 2). Most importantly, the experts' distribution of the estimated cue validities showed a more pronounced dispersion than the novices', as measured by the standard deviation of each participant's validity estimates across the cues: $Ms = 27.5$ ($SD = 7.9$) vs. 19.1 ($SD = 8.4$), $t(69) = 4.28$, $p = .001$. Among the experts, professional experience was unrelated to cue dispersion, $r = .10$, $p = .59$. Confidence in the validity estimates did not differ between experts and novices ($p = .21$) and is therefore not considered further here.

3.2. Did experts and novices differ in strategy selection?

There were clear differences between the experts' and novices' decisions. In 20 of the 45 pair comparisons, the expert and the novice groups differed in terms of the passenger picked by the majority within each group. Moreover, across the 45 pairs of passenger profiles, the experts picked, on average, the same passenger 71.6% ($SD = 13.4$) of the time, whereas for the novices that was the case only 61.7% ($SD = 6.5$) of the time, $t(44) = 4.33$, $p = .001$.

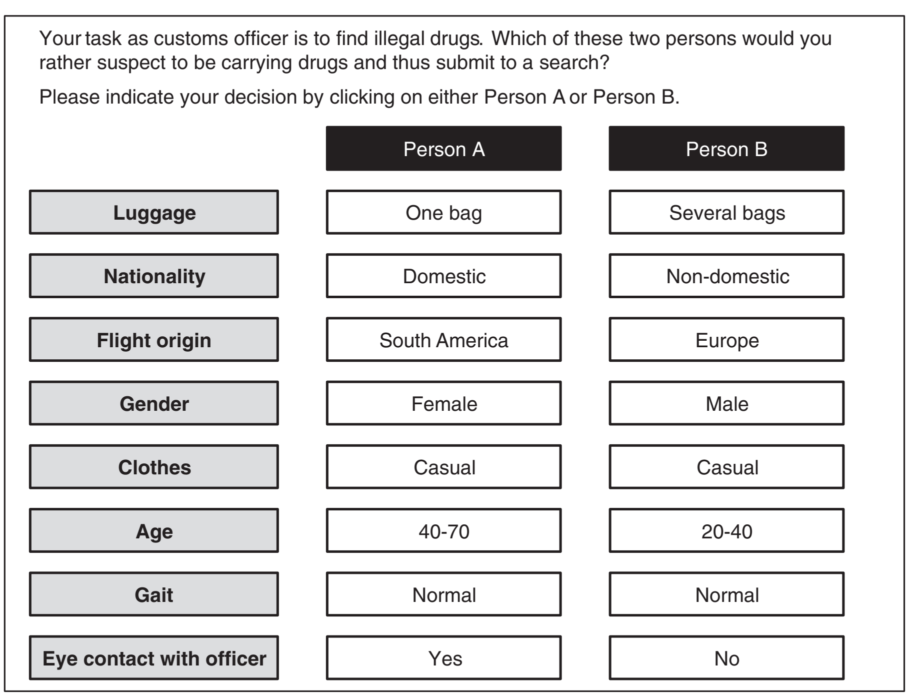

Fig. 1. Screenshot (translated into English) showing how the pairs of passenger profiles were presented to participants in the decision task.

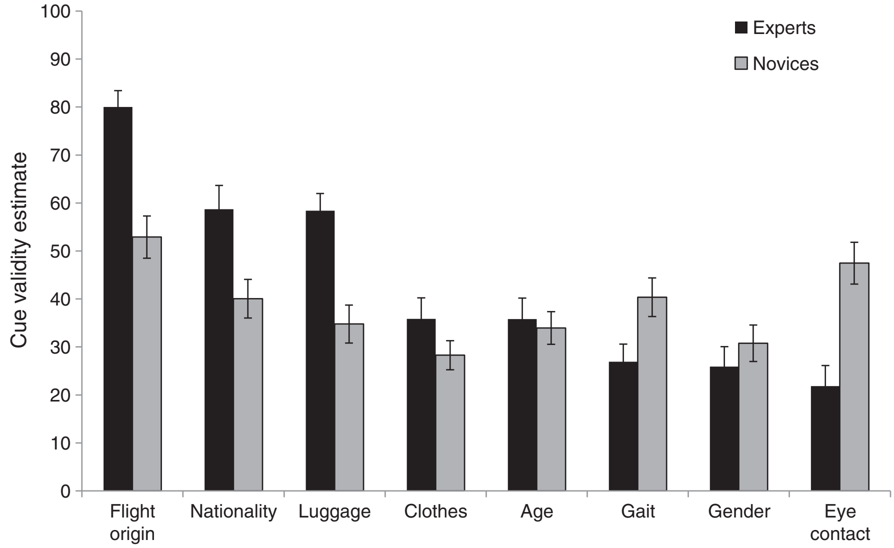

Fig. 2. Experts' (i.e., customs officers') and novices' mean validity estimates of the different cues. Error bars represent ±1 standard error.

To examine the strategies underlying the experts' and the novices' decisions, we first determined, for each individual participant, the decisions predicted by weighted-additive, equal-weight, and take-the-best for the 45 pair comparisons. The predictions were based on the positive and negative cue values (as defined in Table 1 and coded as 1s and 0s, respectively) of the passengers in a pair. As described above, weighted-additive multiplies the cue values by cue weights. As cue weights, we used each participant's subjective validity estimates (as assessed in the cue validity estimation task). For take-the-best's sequential cue inspection, we used each participant's subjective cue hierarchy (as assessed in the ranking task). As the strategies' predictions were based on subjective cue ranks and cue validities, the percentage of critical items (i.e., where two strategies made opposite predictions; recall that the items were selected such that no strategy had to guess) varied across participants. For each participant, there were on average 27.2 ($SD = 9.8$) items where take-the-best and equal-weight made opposite predictions, 14.3 ($SD = 11.0$) items where take-the-best and weighted-additive made opposite predictions, and 11.9 ($SD = 10.0$) items were equal-weight and weighted-additive made opposite predictions.

The predictions for the individual strategies were then compared to participants' decisions using a maximum likelihood approach (see Pachur & Galesic, in press; for a model recovery test of this method, see Pachur & Aebi-Forrer, in press). Accordingly, we determined for each participant i the goodness of fit of strategy k as

$$G^2_{i,k} = -2\sum_{j=1}^{N} \ln\left[f_j(y)\right], \tag{1}$$

where $f_j(y)$ represents the probability with which a strategy predicts an individual decision y on item j. That is, if an observed decision coincides with the strategy's prediction, $f_j(y) = 1 - \varepsilon_{i,k}$; otherwise $f_j(y) = \varepsilon_{i,k}$, where $\varepsilon_{i,k}$ represents participant i's application error (across all N pairs of passengers) for strategy k. For each strategy, $\varepsilon_{i,k}$ was estimated as the proportion of decisions that deviated from the strategy k's prediction (which represents the maximum likelihood estimate of this parameter; see Bröder & Schiffer, 2003). The lower the G^2, the better the model fit. Each participant was classified as following the strategy to which her decisions showed the best fit (i.e., the strategy with the lowest G^2). If the G^2 of the best-fitting strategy equaled or was higher than the G^2 of a guessing strategy (i.e., $\varepsilon = 0.5$), then the participant was classified as guessing.

Which were the strategies that best described the experts' and the novices' decisions? As Fig. 3 shows, the large majority of the experts (64.5%) were classified as following the noncompensatory take-the-best heuristic. The compensatory strategies weighted-additive (14.5%) and equal-weight (17.7%), by contrast, played only a minor role and were significantly less prevalent than take-the-best (weighted-additive: $z = 7.90$, $p = .001$; equal-weight: $z = 6.82$, $p = .001$). As indicated by a significant association between group (experts vs. novices) and strategy (compensatory vs. noncompensatory), the distribution of participants across the compensatory and noncompensatory strategies differed between the expert and the novice groups, χ^2 (1, $N = 65$) = 4.61, $p = .05$ (exact significance).[3] In the latter, the majority of participants (52.9%) were classified as following one of the compensatory strategies (equal-weight: 30.8%; weighted-additive: 22.1%), somewhat more than in the expert group (32.2%), $z = 1.74$, $p = .08$. Compared with the experts, considerably fewer participants in the novice group were classified as following take-the-best, 34.6% ($z = 2.50$, $p = .001$). A slightly larger proportion of novices than of experts was classified as guessing (12.5% vs. 3.2%; $z = 1.39$, $p = .16$).[4] Among the experts, use of take-the-best (coded as: 1 = classified as using take-the-best; 0 = classified as using a compensatory strategy) was weakly associated with professional experience, $r = .25$, $p = .17$. Note, however, that this correlation is potentially restricted due to high homogeneity in strategy use (two-thirds of the experts were classified as following take-the-best).

[3] For this analysis, we collapsed the two compensatory strategies into one category and eliminated the guessing category (to which one and five participants in the expert and novice groups, respectively, were classified) as otherwise too many cells would have expected frequencies smaller than 5, which can reduce the power of the χ^2 analysis.

[4] In additional analyses, we also considered the *take-two heuristic* in the strategy classification (Dieckmann & Rieskamp, 2007). Like take-the-best, take-two is a lexicographic heuristic and searches cues in decreasing order of validity; but unlike take-the-best it stops search only when two cues that favor the same alternative have been found. Therefore, take-two's information processing has both compensatory and noncompensatory aspects. The distribution of strategy users when including take-two in the strategy classification was as follows. Among the experts, there were 64.5%, 8.1%, 9.7%, 14.5%, and 3.2% of participants classified as following take-the-best, take-two, weighted-additive, equal-weight, and guessing, respectively. Among the novices, there were 33.3%, 16.3%, 20.8%, 22.1%, and 7.5% of participants classified as following take-the-best, take-two, weighted-additive, equal-weight, and guessing, respectively. Note, however, that adding take-two to the strategy classification leads to a decreased classification confidence: the Bayes factor for the classifications decreased to 4.15 and 2.47 for the experts and novices, respectively. The resulting distribution should thus be treated with caution.

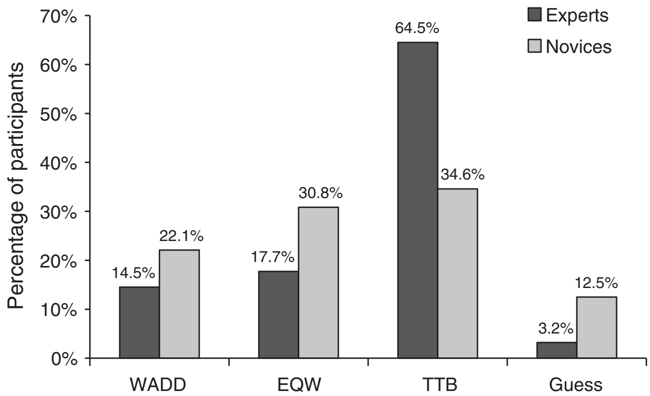

Fig. 3. Classification of experts and novices across the different decision strategies: weighted additive (WADD), equal weight (EQW), take-the-best (TTB), and guessing.

As a measure of classification confidence, we calculated a Bayes factor (BF) for each strategy classification.[5] A Bayes factor in the range of 1 to 3, 3 to 10, and larger than 10 indicates anecdotal, substantial, and strong evidence, respectively, for the classification (Jeffreys, 1961). Across participants (excluding participants classified as guessing), the median Bayes factor was $BF = 4.86$, indicating substantial evidence, and it did not differ between the expert and novice groups, $p = .17$ (as indicated by a median test).

3.3. Does dispersion of the cue validity estimates mediate expert–novice differences in strategy selection?

Given that the experts were more frequently classified as following take-the-best than the novices and that the cue validity estimates of the former showed a higher dispersion, we examined whether cue dispersion might mediate the differences in use of take-the-best between experts and novices. To that end, we conducted a bootstrapping mediation analysis as recommended by Shrout and Bolger (2002). Table 2 shows the results (based on 10,000 runs).[6] Most importantly, although the 95% confidence interval of the weight for the indirect effect from expertise to take-the-best use (i.e., $a \times b$) included zero, there were indications for a partial mediation. Specifically, when dispersion was entered into a logistic regression predicting take-the-best use (using the binary coding described above) based on expertise, this reduced the effect of expertise such that its 95% confidence interval included zero. The P_M ratio, quantifying the strength of mediation (Shrout & Bolger, 2002), was .11, indicating that about a tenth of the effect of expertise on strategy use was mediated by differences in the dispersion of the subjective cue validities. This analysis suggests that cue dispersion might contribute to the expert-novice differences in strategy use but also points to the operation of additional factors.

4. Discussion

Salthouse (1991) proposed that experts might be less constrained than novices in combining and integrating information when making decisions. Our analysis of the decisions of airport customs officers suggests, by contrast, that experts prefer simple decision strategies that rely on few cues and go without integration, whereas novices tend to use compensatory strategies that integrate multiple cues. Among the compensatory strategies, weighted-additive provided the worst account of participants' decisions. This contradicts the claim that much of people's decision making is based on an automatic weighted integration process (Glöckner & Betsch, 2008).

Table 2
Results of the bootstrap mediation analysis of the effect of expertise on the use of take-the-best as mediated by the dispersion of the subjective cue validities.

	Md	95% CI
A	0.42	[0.18, 0.62]
B	0.09	[−0.22, 0.42]
C	0.30	[0.03, 0.54]
c'	0.26	[−0.05, 0.54]
$a \times b$	0.03	[−0.09, 0.20]
P_M	0.11	[−0.51, 1.00]

Note: Path *a* indicates the association between expertise and cue dispersion, path *b* indicates the association between cue dispersion and take-the-best use, and path *c* indicates the association between expertise and take-the-best use. Path *c'* indicates the association between expertise and take-the-best use controlling for cue dispersion. P_M is the ratio of the indirect effect over the total effect of expertise on the use of take-the-best (see Shrout & Bolger, 2002, for details). CI = confidence interval.

Our results represent an important extension of previous findings on expert–novice differences in strategy use (Garcia-Retamero & Dhami, 2009) and also echoes evidence presented by Dhami (2003) that professional judges sometimes rely on simple, noncompensatory decision trees. Experts seem to rely on simple strategies even in a domain in which they make hundreds of decisions every day and obtain regular feedback (although this feedback may be incomplete; we turn to this issue shortly). Moreover, the differences between experts and novices hold when potentially confounding factors such as age and education are properly controlled. Importantly, our results also provide some evidence that the differences in experts' and novices' strategy selection are to some extent adaptive. Relative to the novices', the experts' representation of the cue weight distribution is more skewed; and with a skewed distribution of cue weights, the use of noncompensatory strategies is more appropriate (e.g., Hogarth & Karelaia, 2007). The greater reliance on a compensatory strategies by the novices is consistent with findings from a more artificial inference task by Rieskamp and Otto (2006), asking student participants to judge the creditworthiness of companies; these authors observed that in the absence of experience with a decision domain, people seem to have an initial tendency to use a compensatory strategy, which in principle allows them to explore the task more than does a noncompensatory strategy, that ignores cues (see also Bröder & Schiffer, 2006).

Our study expands on previous research in several ways. First, to our knowledge, it is the first investigation of decision making in the customs domain. Second, it highlights the importance of not only considering the number of cues that expert and novice decision makers use (as is often done in expert–novice comparisons, e.g., Phelps & Shanteau, 1978; Shanteau, 1992), but also formally modeling the strategies used to process these cues. Third, it illustrates how examining the internal representation of the decision environment can reveal the potentially adaptive nature of expert–novice differences in strategy use. Finally, our results demonstrate that simple, noncompensatory heuristics may be an important mental tool for inference even beyond artificially constructed laboratory tasks.

As described above, however, the differences in cue dispersion can only partially account for the differences in strategy use. Other factors shaping the experts' use of simple heuristics may be that customs officers have to make their decisions within a limited time frame and under considerable workload (given the large numbers of passengers generally passing through customs simultaneously). Such a situation fosters reliance on simple noncompensatory strategies (e.g., Pachur & Hertwig, 2006; Rieskamp & Hoffrage, 2008). The officers' use of take-the-best in our study might thus to some extent also reflect a decision routine spilling over from their professional work.

What are the implications of our results for current proposals of the mechanisms underlying strategy selection? The association

[5] The Bayes factor is defined based on the Bayesian Information Criterion (BIC) differences between the best-fitting and the second-best fitting strategy, $BF = \exp(-\frac{1}{2}\Delta BIC)$ (for details, see Wasserman, 2000). The BIC for each strategy is defined as $\text{BIC} = G^2 + k \times \log(n)$, with k being the number of free parameters (which equals 1 for all strategies) and n being the number of decisions.

[6] As the outcome variable was binary, we used logistic regression and therefore estimated the path weights as described by MacKinnon and Dwyer (1993).

between strategy use and cue representation indicates that some of the expert–novice differences may be couched within Marewski and Schooler's (2011) cognitive niches framework, which proposes that strategy selection arises from an interplay between mind and environment. Specifically, the stronger differentiation in the experts' representation of the cue validities (assuming that it stems from an interaction with the environment) may represent a more appropriate "cognitive niche" for the application of a lexicographic strategy such as take-the-best than the novices' less differentiated cue representation. Second, to the extent that the experts' reliance on simple strategies also reflects factors such as limited time or cognitive resources, more explicit processes that trade off a strategy's cost against its expected accuracy, as highlighted by Payne et al. (1993), may also play a role.

It should be highlighted that although experts and novices differed considerably in terms of both strategy use and representation of the cue environment, this does not imply that the experts' decisions (or their cue representations) are more accurate. On the one hand, the customs officers showed greater consensus in terms of both their cue ranking and their decisions. This suggests that they may also show greater individual consistency, which has been linked to accuracy (e.g., Goldberg, 1968, 1970). In addition, compared with the novices' cue ranking, the customs officers' ranking agreed more with the ranking of the chief officer, who according to internal airport statistics had the highest "success" rate (in terms of detected infringements). On the other hand, it is important to note that the officers operate in a "wicked" learning environment (Hogarth, 2001). Specifically, they receive feedback only about the passengers they screen (not about the passengers they do not screen). A further complication is that the screened passengers are likely to represent a skewed sample of the passenger population. This might make it difficult to learn the actual predictive strength of the cues (Dawes, Faust, & Meehl, 1989; but see Elwin, Juslin, Olsson, & Enkvist, 2007). Given these conditions, one cannot exclude the possibility that the stronger consensus among the experts to some extent reflects shared erroneous beliefs about the cue validities. Note that in her study on bail decisions by professional judges, Dhami (2003) found that the simple decision trees on which the judges seemed to rely were partly based on irrelevant information. Such factors might explain why, although expertise is sometimes associated with higher decision quality (e.g., Pachur & Biele, 2007), this is not generally the case (as shown in the meta-analysis by Garb, 1989). In fact, it has been shown that less knowledge can lead to better decisions (Goldstein & Gigerenzer, 2002; Pachur, 2010).

Irrespective of the accuracy of the experts' decisions, an interesting issue for future research concerns the processes by which they learn about the statistical structure of the environment. Additional interviews with the customs officers revealed that they are given no formal instructions on how to proceed in conducting passenger checks and that there are no statistics about the predictive strength of various cues. However, there does seem to be considerable informal exchange of knowledge and subjective experience among officers. Further studies could examine more closely how this knowledge is transmitted. For instance, it is possible that customs officers communicate their experience in the form of simple checklists (e.g., noncompensatory decision trees, which have been discussed as useful and effective decision aids; Katsikopoulos, Pachur, Machery, & Wallin, 2008).

Although there are practically no comparative studies of expert and novice decision making that simultaneously consider both strategy use *and* decision quality (see Ettenson et al., 1987; Garcia-Retamero & Dhami, 2009), our results underline the potential value of such an approach (but note that due to the low base rate of smuggling, it may be difficult to assess decision accuracy in the customs domain). Together with other findings (Garcia-Retamero & Dhami, 2009), our analyses add to the increasing evidence that, in contrast to common belief (e.g., Kahneman, 2011; Salthouse, 1991), intuitive expertise in decision making—at least in some situations—may not reflect the consideration of multiple cues, but the use of simple heuristics.

References

Bröder, A. (2011). The quest for take the best—Insights and outlooks from experimental research. In G. Gigerenzer, R. Hertwig, & T. Pachur (Eds.), *Heuristics: The foundations of adaptive behavior* (pp. 364–382). New York: Oxford University Press.

Bröder, A., & Schiffer, S. (2003). Bayesian strategy assessment in multi-attribute decision research. *Journal of Behavioral Decision Making*, *16*, 193–213.

Bröder, A., & Schiffer, S. (2006). Adaptive flexibility and maladaptive routines in selecting fast and frugal decision strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 904–918.

Camerer, C. F., & Johnson, E. J. (1991). The process-performance paradox in expert judgment: How can experts know so much and predict so badly? In K. A. Ericsson, & J. Smith (Eds.), *Towards a general theory of expertise: Prospects and limits* (pp. 195–217). New York: Cambridge Press.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, *34*, 571–582.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668–1674.

Dhami, M. K. (2003). Psychological models of professional decision-making. *Psychological Science*, *14*, 175–180.

Dieckmann, A., & Rieskamp, J. (2007). The influence of information redundancy on probabilistic inferences. *Memory and Cognition*, *35*, 1801–1813.

Elwin, E., Juslin, P., Olsson, H., & Enkvist, T. (2007). Constructivist coding: Learning from selective feedback. *Psychological Science*, *18*, 105–110.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, *49*, 725–747.

Ettenson, R., Shanteau, J., & Krogstad, J. (1987). Expert judgment: Is more information better? *Psychological Reports*, *60*, 227–238.

European Commission Taxation and Customs Union (2012). Statistics of customs detentions recorded at the external borders of the EU-2010. Retrieved from. http://ec.europa.eu/taxation_customs/customs/customs_controls/counterfeit_piracy/statistics/index_en.htm

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin*, *105*, 387–396.

Garcia-Retamero, R., & Dhami, M. K. (2009). Take-the-best in expert-novice decision strategies for residential burglary. *Psychonomic Bulletin and Review*, *16*, 163–169.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.

Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1055–1075.

Glöckner, A., Heinen, T., Johnson, J., & Raab, M. (2012). Network approaches for expert decisions in sports. *Human Movement Science*, *31*, 318–333.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*, 1–40.

Goldberg, L. R. (1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist*, *23*, 483–496.

Goldberg, L. R. (1970). Man versus model of man. *Psychological Bulletin*, *73*, 422–432.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, *109*, 75–90.

Hogarth, R. M. (2001). *Educating intuition.* Chicago: The University of Chicago Press.

Hogarth, R. M., & Karelaia, N. (2007). Heuristic and linear models of judgment: Matching rules and environments. *Psychological Review*, *114*, 733–758.

Jeffreys, H. (1961). *Theory of probability.* Oxford, UK: Oxford University Press.

Kahneman, D. (2011). *Thinking: Fast and slow.* New York: Farrar, Strauss, Giroux.

Katsikopoulos, K., Pachur, T., Machery, E., & Wallin, A. (2008). From Meehl (1954) to fast and frugal heuristics (and back): New insights into how to bridge the clinical–actuarial divide. *Theory and Psychology*, *18*, 443–464.

Khader, P. H., Pachur, T., Meier, S., Bien, S., Jost, K., & Rösler, F. (2011). Memory-based decision making with heuristics involves increased activation of decision-relevant memory representations. *Journal of Cognitive Neuroscience*, *23*, 3540–3554.

Klein, G. (1998). *Sources of power: How people make decisions.* Cambridge, MA: MIT Press.

Lipshitz, R. (2000). Two cheers for bounded rationality [Commentary]. *The Behavioral and Brain Sciences*, *23*, 756.

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediating effects in prevention studies. *Evaluation Review*, *17*, 144–158.

Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review*, *118*, 393–437.

Mata, R., Pachur, T., von Helversen, B., Hertwig, R., Rieskamp, J., & Schooler, L. J. (2012). Ecological rationality: A framework for understanding and aiding the aging decision maker. *Frontiers in Decision Neuroscience*, *6*, 19.

Mata, R., Schooler, L. J., & Rieskamp, J. (2007). The aging decision maker: Cognitive aging and the adaptive selection of decision strategies. *Psychology and Aging*, *22*, 796–810.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis: University of Minnesota.

Newell, B. R. (2005). Re-visions of rationality. *Trends in Cognitive Sciences*, *9*, 11–15.

Pachur, T. (2010). Recognition-based inference: When is less more in the real world? *Psychonomic Bulletin and Review*, *17*, 589–598.

Pachur, T., & Aebi-Forrer, E. (2013). Selection of decision strategies after conscious and unconscious thought. *Journal of Behavioral Decision Making*, http://dx.doi.org/10.1002/bdm.1780 (in press).

Pachur, T., & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica*, *125*, 99–116.

Pachur, T., & Galesic, M. (2013). Strategy selection in risky choice: The impact of numeracy, affect, and cross-cultural differences. *Journal of Behavioral Decision Making*, http://dx.doi.org/10.1002/bdm.1757 (in press).

Pachur, T., & Hertwig, R. (2006). On the psychology of the recognition heuristic: Retrieval primacy as a key determinant of its use. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 983–1002.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: Cambridge University Press.

Phelps, R. H., & Shanteau, J. (1978). Livestock judges: How much information can an expert use? *Organizational Behavior and Human Performance*, *21*, 209–219.

Rieskamp, J., & Hoffrage, U. (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica*, *127*, 258–276.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology. General*, *135*, 207–236.

Salthouse, T. A. (1991). Expertise as the circumvention of human processing limitations. In K. A. Ericsson, & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 286–300). New York: Cambridge University Press.

Shanteau, J. (1992). How much information does an expert use? Is it relevant? *Acta Psychologica*, *81*, 75–86.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*, 422–445.

Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology*, *44*, 92–107.